Remote Sensing in Ecology and Conservation Open Access

ZSL LET'S WORK FOR WILDLIFE

ORIGINAL RESEARCH

Mapping the world's coral reefs using a global multiscale earth observation framework

Mitchell B. Lyons[1], Chris M. Roelfsema[1], Emma V. Kennedy[1], Eva M. Kovacs[1], Rodney Borrego-Acevedo[1], Kathryn Markey[1], Meredith Roe[1], Doddy M. Yuwono[1], Daniel L. Harris[1], Stuart R. Phinn[1], Gregory P. Asner[2], Jiwei Li[2], David E. Knapp[2], Nicholas S. Fabina[2], Kirk Larsen[3], Dimosthenis Traganos[1] & Nicholas J. Murray[1,4]

[1]Remote Sensing Research Centre, School of Earth & Environmental Sciences, The University of Queensland, Queensland 4072, Australia
[2] Center for Global Discovery and Conservation Science, Arizona State University, Arizona 85287, USA
[3] Vulcan Inc, Washington 98104, USA
[4]College of Science & Engineering, James Cook University, Queensland 4811, Australia

Keywords
Coral reefs, machine learning, mapping, monitoring, remote sensing

Correspondence
Mitchell B. Lyons, Remote Sensing Research Centre, School of Earth & Environmental Sciences, The University of Queensland, Queensland 4072, Australia. Tel: +61 7 3365 6455; Fax: +61 7 3365 6899; E-mail: mitchell.lyons@gmail.com

Editor: Nathalie Pettorelli
Associate Editor: Vincent Lecours

Received: 20 November 2019; Revised: 11 February 2020; Accepted: 28 February 2020

doi: 10.1002/rse2.157

Abstract

Coral reefs are among the most diverse and iconic ecosystems on Earth, but a range of anthropogenic pressures are threatening their persistence. Owing to their remoteness, broad spatial coverage and cross-jurisdictional locations, there are no high-resolution remotely sensed maps available at the global scale. Here we present a framework that is capable of mapping coral reef habitats from individual reefs (~200 km^2) to entire barrier reef systems (200 000 km^2) and across vast ocean extents (>6 000 000 km^2). This is the first time this has been demonstrated using a consistent and transparent remote sensing mapping framework. The ten maps that we present achieved good accuracy (78% mean overall accuracy) from multiple input image datasets and training data sources, and our framework was shown to be adaptable to either benthic or geomorphic reef features and across diverse coral reef environments. These new generation high-resolution map data will be useful for supporting ecosystem risk assessments, detecting change in ecosystem dynamics and targeting efforts to monitor local-scale changes in coral cover and reef health.

Introduction

Over the past few decades, remote sensing has become a central tool for environmental monitoring and decision-making (Nagendra et al. 2013; Rose et al. 2015). In particular, remote sensing can deliver important environmental records that allow consistent monitoring of variables such as vegetation health and production, and ocean colour and temperature. There is now a renewed focus on frameworks to deliver near real-time information on thematic biophysical attributes like land cover and habitat types (Foo and Asner 2019). This explicitly links to the growing importance of ecosystem modelling and ecosystem risk assessments for understanding the health, status and trajectory of ecosystems (Murray et al. 2018a).

Maps continue to provide a foundational basis for grounding ecological monitoring and modelling over space and time. At regional to global scales, satellite earth observation data offer the only viable source of information suitable for mapping and monitoring ecosystems (Hansen et al. 2010). As earth observation has matured as a field, so has the array of information types derived from sensor data (Nagendra et al. 2013). We are transitioning towards being able to provide continuous data about natural resources, for example vegetation cover and height, water depth and chlorophyll content (Coops and Wulder 2019). Nevertheless, thematic habitat maps that depict discrete cover classes remain the primary data source used in many legislative frameworks, monitoring programs and scientific applications. Indeed, habitat mapping has been identified as a key technology for coral reef conservation and restoration (Foo and Asner 2019; Purkis et al. 2019).

There are now multiple examples of remote sensing frameworks that are implemented at large spatial extents and at high resolution (e.g. <30 m pixels) (Hansen et al. 2010; Murray et al. 2019). They have largely shown that some of the traditional limitations – spatial extent, timely re-analysis, user-friendly methods – have been overcome. Thus, focus is now moving towards developing methods that provide useful information at multiple spatial scales (Nagendra et al. 2013). These methods should require little modification to handle new observations, such as incorporating new satellite data sources or observations over time, or to deliver different mapping outputs, such as a transition from thematic to continuous variables (Coops and Wulder 2019). Methods should also have the potential to be implemented with minimal technical skills to facilitate better blending between remote sensing specialists and practitioners and promote use as a monitoring tool.

Future mapping frameworks should therefore not only be able to generate detailed habitat maps over large spatial extents, but they should strive for attributes that enable their use in management and conservation monitoring systems (Nagendra et al. 2013; Rose et al. 2015; Murray et al. 2018a; Coops and Wulder 2019; Foo and Asner 2019; Stehman and Foody 2019). Key attributes for such frameworks are: (1) ability to access and process multisource sensor data within a single analysis platform; (2) flexibility to incorporate training data from a range of sources; (3) capacity to provide thematically accurate data from local (e.g. an individual reef) to global (e.g. entire reef system) scales; (4) ability to update outputs when new sensor or training data are made available; (5) capacity for modifying the map output type; (6) implementation on a publicly available analysis platform that requires minimal local computing resources; and (7) achieve accuracies that meet the expectations of ecosystem managers. Most of these key attributes (particularly 4 and 5) require a mapping framework that is easy, efficient and timely to 're-run', by the original investigators or by different individuals or groups. Recent large-scale mapping efforts are fulfilling these requirements (Murray et al. 2019), but a notable exception is coral reef environments. Like many global ecosystems, coral reefs face imminent threats, but large-scale mapping efforts, while significant, have not been developed via methodologies that can be repeated in a consistent or timely manner.

Reef mapping has a long history, from Charles Darwin's early global distribution estimates, to massive extents of semi-automated classification of high-resolution satellite imagery. The first 'modern' attempt to catalogue the world's coral reefs in a spatially explicit manner was the United Nations Environment Program World Conservation Monitoring Centre (UNEP-WCMC) coral reef mapping project (Spalding et al. 2001). The global map compiled data from a variety of sources, ranging from navigational charts to the (Sheppard and Wells 1988) collections of individual reef maps, with variable and often incompatible classification schemes.

Currently, the majority of the UNEP-WCMC coral reef map comprises maps originating from the Millennium Coral Reef Mapping project (Andrefouet et al. 2006). These maps were significantly different to past mapping efforts in that a consistent classification scheme was used, along with a consistent satellite image data source (Landsat 7, 30 m pixels). Despite these global layers providing critical information on reef distributions for nearly 20 years, there is a desire for coral reef maps at a higher spatial resolution, and for maps that provide information on both geomorphic zonation and benthic habitat type. There is also a need to reduce the amount of manual image interpretation required, upon which much of the Millennium project was based.

A recent example that solved these challenges over a very large extent used a semi-automated remote sensing method, combining high-resolution satellite imagery, object-based image analysis and *in situ* field data (Purkis et al. 2019). Despite the large mapping extents, it did not map entire reef systems, it is not amenable to timely re-implementation (new input data sources, different classification schemes) and the methods are not accessible to users from all socio-economic backgrounds.

Here we present a mapping framework that achieves the desirable attributes listed above to underpin future coral reef mapping efforts at local to global scales. We demonstrate the framework by progressively mapping an individual reef (~200 km^2 of reef) to the entire Great Barrier Reef (~200 000 km^2 of reef) and, finally, testing the method across the South West Pacific coral reef region (~140 000 km^2 of reef). This region contains highly complex reef environments distributed widely across 6 000 000 km^2 of ocean (every reef around and between New Caledonia, Vanuatu, Tuvalu, Tokalau, Samoa, Nuie, Tonga and Fiji). These are both the geographically largest and most detailed coral reef maps derived from a single, consistent and repeatable earth observation analytical approach.

We demonstrate how the framework works, and how it is being adopted by organizations to support coral reef conservation and monitoring efforts globally. The purpose of this paper is not to present a new catalogue of the world's coral reefs, or even an explicit comparison of maps from our new framework to existing maps. Rather this paper presents the foundations to develop a novel coral reef monitoring system that is agile and dynamic enough to support rapidly changing needs, data sources and end users into the future.

Materials and Methods

Case study locations and data

To demonstrate the multiscale nature of our mapping framework, we first use three focus extents on the Great Barrier Reef: individual reef, reef management region

(247 reefs) and entire shelf barrier reef system (~3000 reefs). We illustrate the multimodal nature of the framework using different satellite image data and different types of training data. We then show its ability to transfer to a new environment, mapping the South West Pacific region (>2000 reefs across 6 000 000 km^2), a particularly morphologically complex region with a mix of atolls, barrier reefs and diverse fringing reefs. Each focus extent utilized different combinations of input satellite imagery, bathymetry and wave data with spatial resolution varying from 2 m to 30 m pixel size. Table 1 outlines the various combinations of input data across the focus extents, how much of the focus extent is represented by training and validation data, including references that detail their provenance, pre-processing and analysis methods.

Mapping framework

Our mapping framework combines image segmentation, machine learning prediction and object-based classification into a single, flexible classification approach designed to freely move between focus extents and data types, while simultaneously handling redundant data and wide variation in data quality. The framework has four central processing modules that are applied after selecting a combination of satellite image, bathymetry and wave data, and acquiring reference data to train the classifiers (Fig. 1).

The entire workflow is implemented on Google Earth Engine (Gorelick et al. 2017). Earth Engine is a platform that allows both visualization and analysis of various geospatial datasets, similar to traditional remote sensing or GIS workflows, except implemented in a scalable, cloud-based computing environment. It has a Python and JavaScript application programing interfaces (API), and can be used on any computer with an internet connection. Users can access an existing catalogue of publicly available geospatial datasets, or upload their own data, while maintaining ownership of all algorithms and results. Earth Engine is free for research, education and non-profit use.

Table 1. Mapping area details from the individual reef scale to the whole Great Barrier Reef and South West Pacific region, including reef area mapped, input data combinations and training/validation data type and coverage

Focus extent	[†]Number of reefs, shallow reef area and reef habitat area	Satellite image data (pixel resolution)	Bathymetry data	Wave data	Training and validation inputs (% geographic extent covered)
Heron Reef	1 ~43 km^2 ~220 km^2	Worldview-2 (2 m)[¶¶]	CASI derived (4 m)[‡‡]	Bathymetry/fetch model (Harris)[¶¶]	• Point-based in situ data[9] • Map-derived sub-sampling[10](extent: 100%)
Cairns-to-Cooktown management region	247 ~1,700 km^2 ~3 000 km^2	Planet Dove (5 m) [¶¶]Landsat-8 (15 m)[¶]	Planet Dove derived (5 m)[¶¶] Landsat-8 derived (15 m)[¶]	Bathymetry/fetch model (harris)[¶¶] Bathymetry, wind-gen, propagation model (Callaghan)[10]	Point-based in situ data[9] Map-derived sub-sampling[10] (extent: ~15%)
Whole Great Barrier Reef	~3,000 ~16 000 km^2 ~200 000 km^2	Sentinel-2 (10 m)[††]	Deep Reef Explorer data (30 m)[§§]	Bathymetry, wind-gen, propagation model (Callaghan)[10]	Map-derived sub-sampling[10] (extent: ~1%)
South West Pacific	~ N/A ~16 000 km^2 ~140 000 km^2	Planet Dove (5 m)[¶¶]	Planet Dove derived (5 m)[¶¶]	No wave data	Map-derived sub-sampling[10] (extent: ~5%)

[*]Shallow reef area = reef classes mapped ~<5 m deep; reef habitat area = reef zone mapped < 25 m deep; number of reefs for South West Pacific not estimated due to massive extents of fringing reefs.

[†]Worldview-2 image acquired 1 October 2014, geo-corrected to < 1 m, atmospheric correction via FLAASH.

[‡]Landsat-8 mosaic built from imagery between 2013 and 2016, bathymetry derived via physics-based inversion (Roelfsema et al. 2018).

[§]Planet Dove mosaic built from imagery between 2018 and 2019, bathymetry derived from a ratio-based empirical algorithm (Roelfsema et al. 2018; Li et al. 2019).

[¶]Sentinel-2 mosaic built from median reflectance between 2015 and 2018 on Google Earth Engine (Gorelick et al. 2017).

[††]CASI hyperspectral data acquired 2 July 2002, bathymetry derived via adaptive lookup approach (Hedley et al. 2009; Roelfsema et al. 2018).

[‡‡]*Deep Reef Explorer* dataset - bathymetry compiled from multiple sources (Beaman 2010).

[§§]*SWAN* wave propagation-based wave model that incorporates bathymetry and meteorological data (Callaghan et al. 2015).

[¶¶]*ReefWave* bathymetry and fetch-based wave model (*unpublished*).

[9]In situ data derived via analysed georeferenced photo-transect data; from and derived similarly as (Roelfsema et al. 2018).

[10]Detailed habitat maps created via an object-based analysis routine from or similar (Roelfsema et al. 2018).

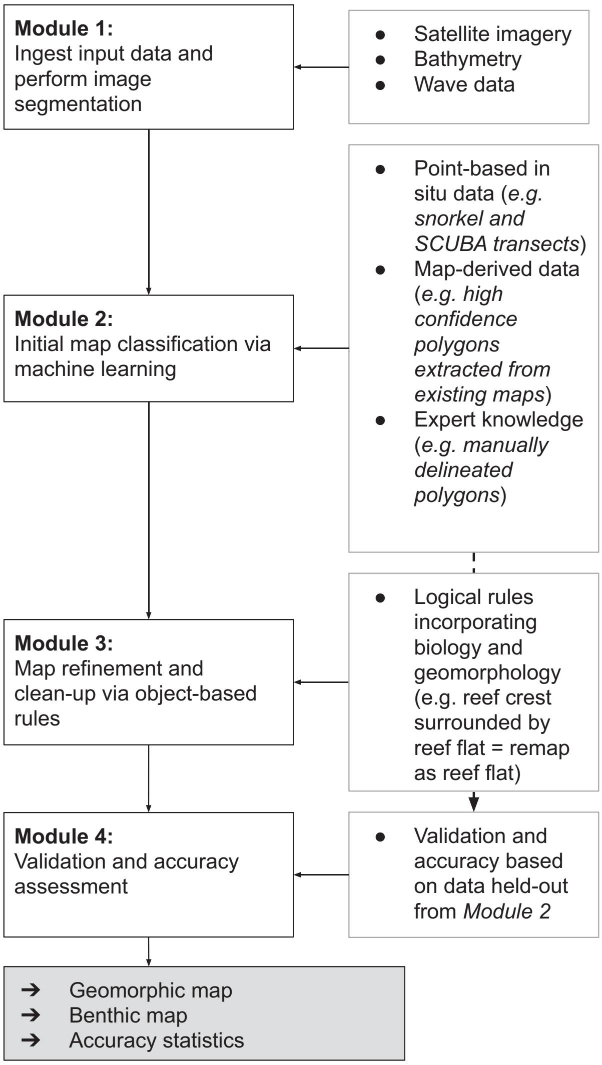

Figure 1. Flow chart detailing the four key modules of the coral reef mapping framework, including the data input types, processing steps and output products.

Data collation and segmentation

First, the relevant covariate data layers and training data are collated into a common geographic reference system. The framework uses both spectral reflectance data and biophysical data that have well-established links to the drivers of distributions of reef biota (Roelfsema et al. 2018). We used reflectance data from Landsat-8, Sentinel-2, Planet Dove and Worldview-2 satellites (Table 1), but most reflectance data could be substituted. Bathymetry, slope angle (derived from bathymetry) and wave data can be used to differentiate most reef geomorphological zones, and are also useful as surrogates for aspects of the physical environment (light availability, temperature, energy) that influence coral reef ecological partitioning. Texture metrics (gray-level co-occurrence matrix; kernel-based neighbourhood variance) are also used to aid detection of other physical and biological aspects of coral reefs, such as surface roughness and habitat heterogeneity.

After the covariate layers are collated, the mapping area is segmented into image 'objects', such that an object represents a relatively homogenous group of pixels, that balances both shape and colour (Blaschke et al. 2014). We used *simple non-iterative clustering* (Achanta and Susstrunk 2017) on the blue, green and depth layers, but other combinations of segmentation algorithms and input data could be used. Briefly, this segmentation method starts with a uniform grid of pixels (centroids), and clusters are formed around these centroids via distance calculation in a *n*-dimensional space of colour and spatial coordinates. Object size is a user-defined parameter that can be easily modified to suit the imagery or required output. Object 'compactness' is also a user-defined parameter that controls the possible geometry, namely roundness and linearity, of an object. This parameter often required trial and error, and we set it to a value (the same for all maps) that allows both round objects as well as linear objects, to account linear features like reef crests and spur and groove formations.

For each object, the mean value for every input data layer was calculated. The pixel-based and object-based data were stacked together such that each pixel location in the stack of covariate layers included both the exact pixel-based value and the values for the object that pixel belongs to. Reefs are complex connected environments, so our hybrid approach is useful because it allows both individual pixel values and neighbourhood information to simultaneously inform the classification.

Coral reef habitat classification

The reference data used for mapping training and validation were from two key sources: (1) point-based field data derived from georeferenced *in situ* photographs from SCUBA and snorkel transects (Roelfsema et al. 2018) or (2) derived by sub-sampling points from high-confidence polygons within existing maps or from expert-derived polygons (image interpretation; Stehman and Foody 2019; Murray et al. 2019). The two rules for selecting training and validation data were: (i) for point-based training data, 75% was used for training and 25% reserved for validation; and (ii) for map-based training data, 3000 random points sampled within each mapping category, 50% for training and 50% for validation. Reference data from which training and validation data were distributed reasonably evenly across the focus extents, although large areas did not have any reference data (Figs. 2 & 3).

The classification module samples the covariate data at locations defined by training data, which represent known occurrences of bottom type, to train a random forest algorithm (James et al. 2013). Random forests are an ensemble decision tree method, where multiple decision trees are created, and (for classification problems) the final prediction is the mode of the prediction from all trees. Each tree contains a random sample (bootstrapped) of the training data and at each node split a randomly selected set of the covariate features is used. The random selection of training data and covariate data ensures uncorrelated decision trees, meaning the random forest method is less prone to overfitting and is robust to redundant covariate data. The random forest classifier was trained with 50 trees per class, a minimum leaf population of 1 and the square root of the total number of covariates as the number chosen at each node split (James et al. 2013).

Once trained, the random forest model is used to predict the class membership of each pixel across the whole focus extent. Here we developed two thematic map types: geomorphological zones (e.g. lagoon, reef flat, reef crest, reef slope) describe the natural structural reef features that underpin the most important biology; and benthic composition (e.g. algae, rubble, coral) describes reef substrates and benthos. These two thematic map outputs are the most commonly mapped thematic structures in coral reef ecosystems (Roelfsema et al. 2018). A full list of classes and their description is in Appendix S1. In this study we have limited the mapping to depths of around

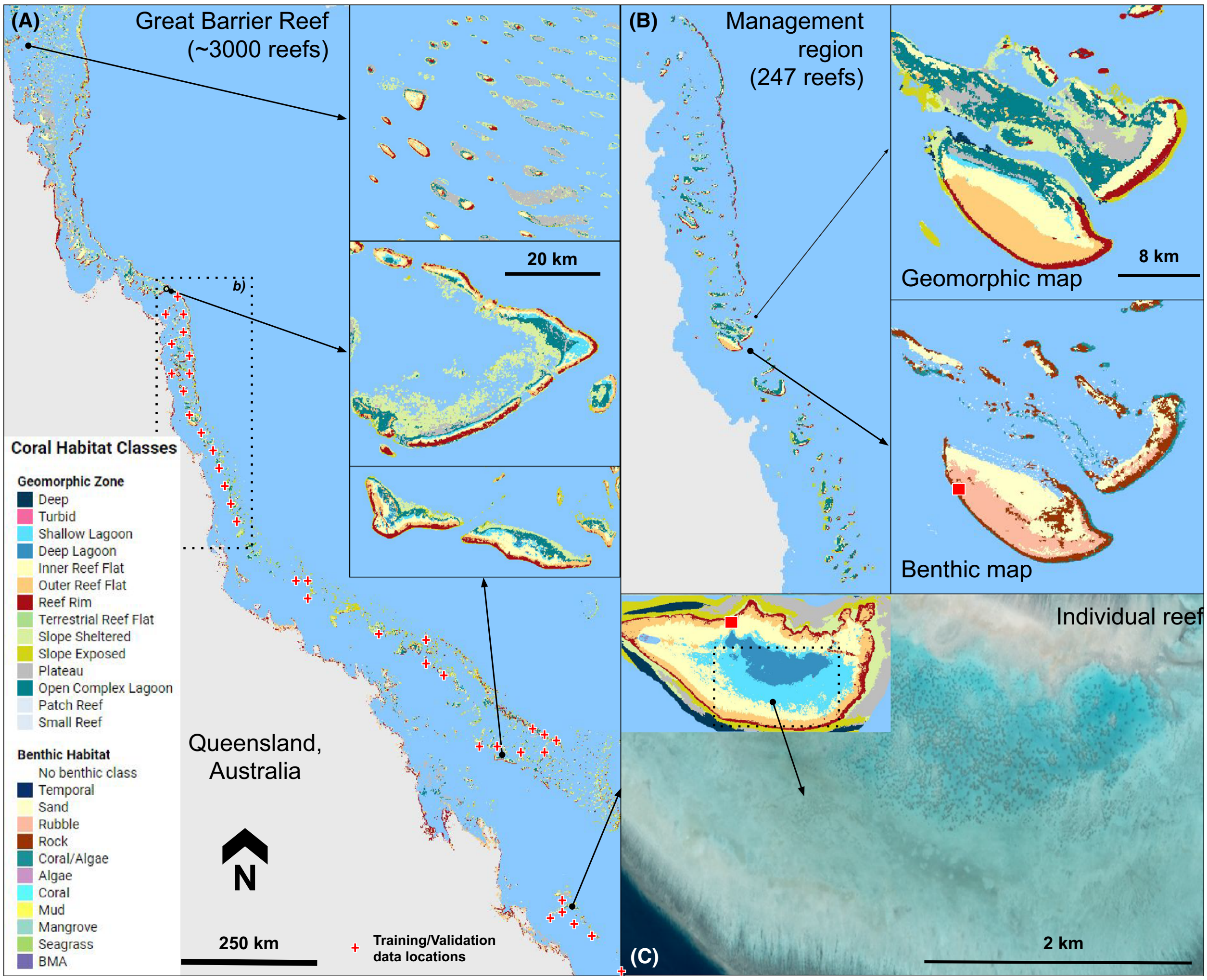

Figure 2. A demonstration of the varying spatial scale and detail possible from the coral reef mapping framework: (A) The Great Barrier Reef (Sentinel-2, 10 m); (B) Cairns-to-Cooktown region (Planet Dove, 5 m); (C) Heron Reef (Worldview-2, 2 m). Red plus symbols indicate training/validation data locations. Small red squares in panel b/c denote zoom location for Figure 4. The maps can be explored in detail here: mitchest.users.earthengine.app/view/coral-map-explorer.

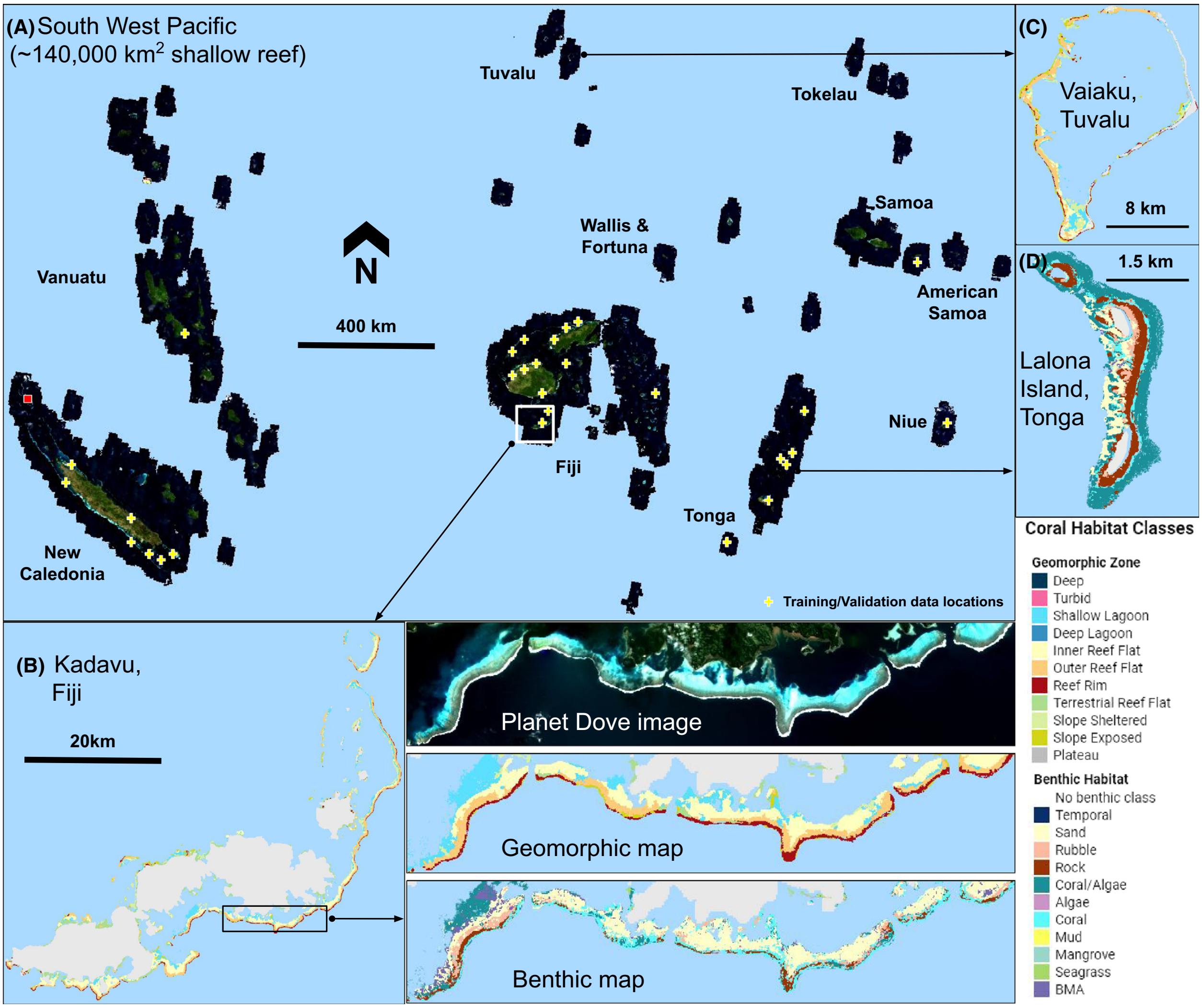

Figure 3. A demonstration of the varying spatial scale and detail possible from the coral reef mapping framework (all Planet Dove 5 m): (A) the South West Pacific region showing the Planet Dove image mosaic; (B) the image mosaic, geomorphic map and benthic map for Kadavu, Fiji; (C) Geomorphic map for Vaiaku, Tuvalu; (D) benthic map for Lolona Island, Tonga. Yellow plus symbols indicate training/validation data locations. Small red square (above New Caledonia) in panel (A) denotes Surprise and Merite reef location for Figure 5. The maps can be explored in detail here: mitchest.users.earthengine.app/view/coral-map-explorer.

15 m due to the water penetration potential of the satellites we used (Li et al. 2019), but the framework is flexible to accommodate different thresholds. Producing each map type (benthic and geomorphic) requires a different set of covariate data layers and training data (Table 1).

The clean-up module applies 'object-based rules', a type of expert system (Pekel et al. 2016), which enables direct translation of geomorphological, ecological and biological principles into logical mapping and contextual algorithms that reduce misclassifications in map outputs. Most of these rules are well defined from previous research on coral reef mapping (Roelfsema et al. 2018). Geomorphic structure is more amenable to logical neighbourhood rules than benthic habitat, as benthic type can be more dynamic in terms of its neighbouring benthic types. For example, consider a small group of pixels (an object) mapped as *reef crest* that are surrounded by *reef flat* pixels – *reef crest* by definition must occur along the edge of reef flat, not surrounded by it, so a class-logic rule is applied to reclassify that group of *reef crest* pixels to *reef flat*. Some classes need to be a minimum size to justify their assignment, for example, lagoons are typically $> 50 \times 50$ m, so very small areas mapped as *shallow lagoon* would be reclassified. Rules for re-classifying benthic classes use the underlying geomorphic classification for the logical rule. For example, *seagrass* is very unlikely to occur on a reef slope, so a class-logic rule reassigns those areas of *seagrass* (spectrally dark) to *coral/algae*, which are the more likely spectrally dark substrates to occur on reef slopes.

A series of these rules, referred to as a 'ruleset' in object-based approaches, were developed for this module. The exact combination of rules varies per mapping area, due to different combinations of classes and reefs type. Typically, about 20 to 30 rules were applied to the geomorphic map, and between 5 and 10 rules were subsequently applied to the benthic map. The full ruleset for this clean-up module is available in our code (see *Code and data access*), which includes a plain English description of each rule.

Accuracy assessment

The final module computes standard accuracy assessment metrics for the output maps. Accuracy is estimated using the data held-out from the classifier (either 25% or 50% of the entire sample depending on data source, see *methods*). To support decision-making, the module provides both a traditional error matrix approach to calculate overall, user and producer accuracy, along with a 95% confidence interval on overall accuracy using a non-parametric bootstrap (Lyons et al. 2018). Accuracy statistics were calculated for each mapping scenario, for both the geomorphic and benthic maps (10 maps in total).

Code and data access

The code for the entire Google Earth Engine framework is accessible online, as both a live version (github.com/CoralMapping/gee-mapping-source) and as a static release as per this paper (https://doi.org/10.5281/zenodo.3714181). Landsat/Sentinel imagery, training data and all bathymetry/wave data are provided open access (Planet and Worldview image data are proprietary).

Results

Our framework successfully generated geomorphic and benthic zone habitat maps from five different earth observation sensors and a range of training data sources, at spatial resolutions between 2 m and 15 m, and at spatial scales from 200 km^2 to 200 000 km^2 (Table 1, Figs. 2–4). The methods transferred seamlessly between the shallow continental shelf reef system of the Great Barrier Reef (predominantly platform reefs) to the complex patchwork of oceanic reef types (mixed fringing reefs, barriers, subtidal atolls and almost atolls) developed across the South West Pacific region (every reef across > 6 000 000 km^2 of ocean around and between New Caledonia, Vanuatu, Tuvalu, Samoa, Nuie, Tonga and Fiji). For the maps that used bathymetry derived from Planet Dove data, occasionally, small reef areas were unable to be mapped when water depth values could not be derived due to data quality and water quality interactions.

In total 10 maps were produced. Overall accuracy of the maps was consistently high (mean 78%; median 80%) and all but one map achieved an overall accuracy of > 70%. There were no notable differences between the geomorphic and benthic maps in terms of accuracy. These accuracies are similar to existing coral reef maps reported in the literature (Roelfsema et al. 2018; Purkis et al. 2019). Full results for the accuracy assessment, including all accuracy measures, error matrices, bootstrapped confidence intervals and individual class accuracies are provided in Appendix S2.

The spatial extent and detail of these maps is difficult to explore and appreciate in static form, therefore we provide a web application built on Google Earth Engine for readers to explore the mapping products (mitchest.users.earthengine.app/view/coral-map-explorer). The high-spatial resolution Planet satellite map products for the Cairns-to-Cooktown and South West Pacific regions can be downloaded from the Allen Coral Atlas (allencoralatlas.org/), and the Great Barrier Reef maps will soon be published by the Great Barrier Reef Marine Park Authority (gbrmpa.gov.au/).

There was broad agreement between the maps from this study and existing maps, although similar to the Millennium (Andrefouet et al. 2006) and Purkis et al. (2019) maps, our maps make a clear information improvement over the UNEP-WCMC map (Fig. 5). At the broad spatial scale, the pattern and zonation was similar between our maps and the Millennium (Andrefouet et al. 2006) and (Purkis et al. 2019) maps (Fig. 5). Inspecting the maps at finer spatial resolution showed that our maps increased the information content over the Millennium (Andrefouet et al. 2006) maps, but were less detailed than the Purkis et al. (2019) maps (Fig. 5). Importantly, our framework provides a separate, coincident geomorphic and benthic map, while the Millennium (Andrefouet et al. 2006) maps only give geomorphic information, and the Purkis et al. (2019) maps give combined geomorphic-benthic habitat information.

Discussion and conclusions

This study demonstrates a mapping framework that is capable of providing varying levels of thematic information detail, while handling vast amounts of data at local to continental scales within a publicly available system that has minimal processing limitations. As the test case, we demonstrated coral reef habitat mapping at spatial

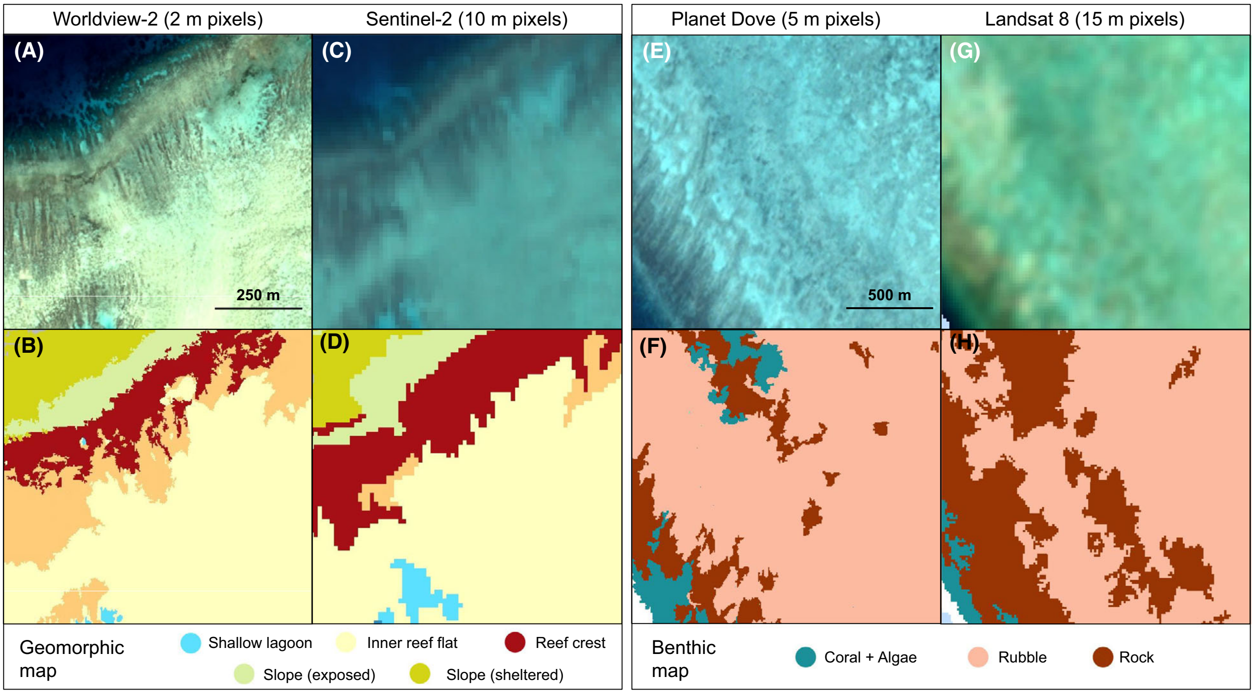

Figure 4. Example coral reef habitat maps on the Great Barrier Reef showing differences in spatial resolution from different satellite sensors for a geomorphic map of Heron Reef (panels A–D; red square in Figure 2 panel C) and a benthic map of Batt Reef (panels E–H; red square in Figure 2 panel B). The maps can be explored in detail here: mitchest.users.earthengine.app/view/coral-map-explorer.

resolution ranging from 2 m to 15 m, from an individual reef (~200 km^2) up to the first ever detailed maps of geomorphic type and benthic habitat for the entire Great Barrier Reef (~200 000 km^2) and South West Pacific region (~140 000 km^2). Our framework is an example of the growing momentum to implement methods capable of near real-time thematic mapping for environmental monitoring (Coops and Wulder 2019; Foo and Asner 2019), and a shift in the remote sensing discipline towards providing methods that facilitate both top-down and bottom-up remote sensing for science and management activities (Murray et al. 2018b).

Coral reef ecosystem risk assessment and monitoring

There is a need for structured ecosystem risk assessments to identify ecosystems at risk of large, detrimental changes, and earth observation has become a critical data source to support these assessments (Murray et al. 2018a). Remote sensing is crucial for understanding the distribution and change of ecosystems, particularly for remote and broadly distributed ecosystems like coral reefs, of which a high proportion (40%) is considered remote and isolated (Foo and Asner 2019; Purkis et al. 2019). Ecosystem risk assessments, as well as assessments of progress towards global conservation targets, require detection and quantification of change over time (Keith et al. 2013), and doing this via a broad un-targeted approach for coral reef environments is cost-prohibitive (Foo and Asner 2019).

High variability and availability of data sources relevant to ecosystem dynamics and varying requirements of output map products (Nagendra et al. 2013; Rose et al. 2015; Coops and Wulder 2019) have also stifled ecosystem risk assessments in coral reef environments (Bland et al. 2017). Effective ecosystem risk assessment thus requires spatially explicit data representing multiple levels of biological organization to support estimates of area change, ecosystem degradation, collapse thresholds and spatially explicit simulation models. In this context, the ability of our framework to deliver coincident geomorphic zonation and benthic habitat maps is particularly advantageous. The growing use of end-to-end ecosystem models (e.g. eReefs; Baird et al. 2018) relies heavily on spatially explicit data, often remote sensing-derived maps (Bland et al. 2017), reinforcing the importance of flexibility in our framework.

In context with other mapping efforts

Tremendous human effort has gone into mapping the planet's coral reefs over the last 200 years – from

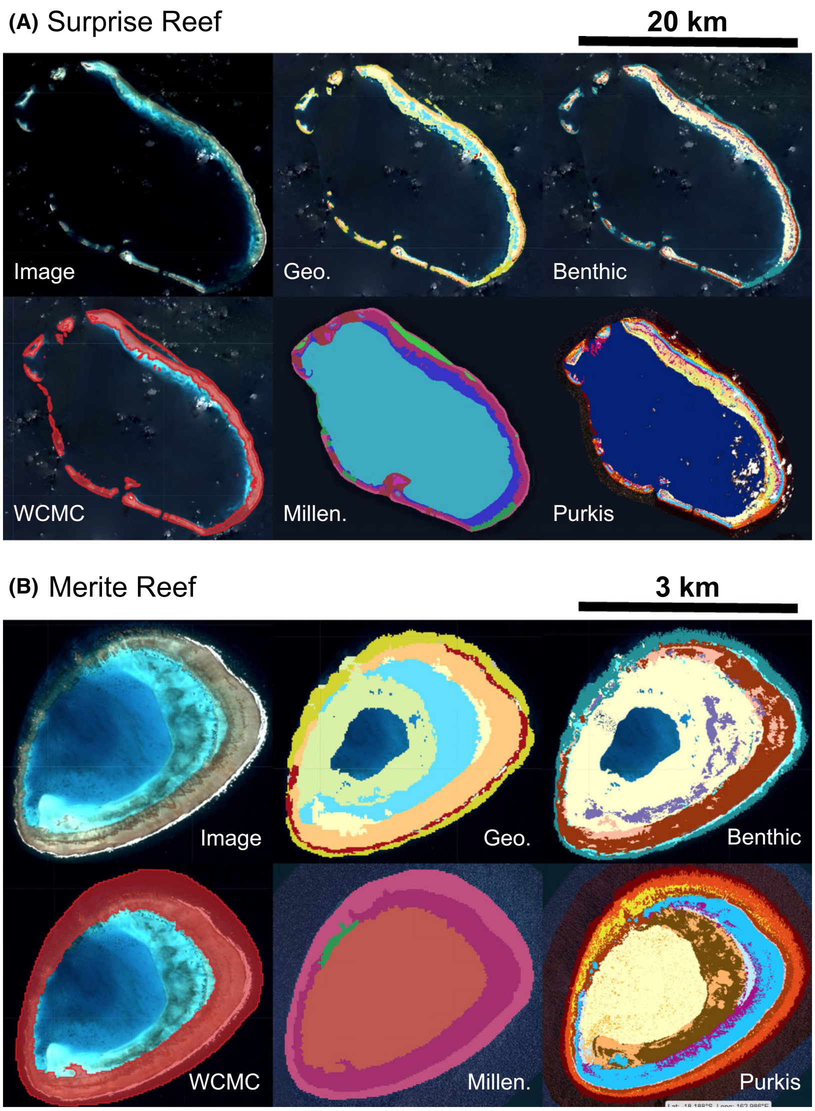

Figure 5. A comparison of the geomorphic and benthic maps produced in this study (*Geo.* & *Benthic*) to the UNEP-WCMC (Spalding et al. 2001), Millennium Project (Andrefouet et al. 2006) and (Purkis et al. 2019) reef maps. Surprise Reef and Merite reef are situated at the north west tip of New Caledonia. This figure provides a broad comparison of class resolution and distribution rather than direct comparison of habitat classification; no class colour definitions are provided as there are around 40 different classes between the mapping projects.

dangerous voyages to discover and chart reefs at the end of the 19th Century, through to the collation and digitization of regional, national and global maps following the widespread availability of computers, satellite imagery and modern mapping methods. The aim of this study is not to compare our outputs with those global-scale coral reef maps in current circulation, but rather to present a new framework that will support and enhance reef mapping efforts worldwide. Future work will focus on refining maps through quantitative analysis, incorporating new training data and comparing with currently in use products. Our initial comparisons have shown broad agreement in terms of both accuracy (see Results section) and broad spatial patterns (Fig. 5), which is encouraging given the scale and detail attainable from our framework.

We expect our maps to support specific localized use where appropriate, and amalgamation into existing resources to facilitate wall-to-wall coverage for existing large-scale management and conservation frameworks. In that sense, mapping coral reefs using consistent methods over such large extents will help facilitate hierarchical and generalized classification schemes for both geomorphic zonation and benthic habitat types, a major concurrent focus of our work.

Key future challenges

A key limiting factor in this framework is the quantity and accuracy of training data and its distribution across space and time. Our framework enables an unprecedented ability to predict beyond the bounds of a focus extent. However, model performance can degrade significantly when predicting beyond the region from which the map classification is trained, and therefore scaling the framework to the global domain requires considerable amounts of training data distributed across the world. Many of the issues relating to training data are explicitly linked to validation data, because scaling up to the global domain requires that large areas have no validation data. Here we infer our reported accuracy into those areas, but it is actually quite uncertain. We encourage users to conduct additional accuracy assessments at their desired scale and provide feedback.

Thus, a key focus of the framework is continued development and investment into the collection and curation of *in situ* and expert-derived data to serve as training and validation data. Our study has shown that, despite considerable advances in mapping technologies, mappers and end users must still continue to scale their investment in training data according to issues such as map relevance (validation in context of class importance/composition), statistical rigour and reproducibility (Stehman and Foody 2019).

Our framework is not dependent on training and validation data of one particular source. It is able to utilize point-based *in situ* data, expert knowledge-derived points or polygons, or high-confidence polygons extracted from existing maps – either as a single source type, or mixed stream of those options. The data sources can also be mixed or separated between training and validation. This is an ideal situation for future efforts to incorporate the growing scientific use of citizen science data (Callaghan et al. 2019) and enhanced collection methods or existing archives (e.g. coralwatch.org, hawaiicoral.org). Flexibility in training and validation inputs will allow citizen science data to enhance user feedback mechanisms and increase the pool of training and validation data. Flexible data structure will also help the transition towards more user-friendly implementations of mapping methods, such as online tools that allow users to make habitat maps by inputting their own training and validation data (Murray et al. 2018b).

Continued investment into mapping

An agile mapping framework such as ours offers an opportunity for investment from management agencies or conservation organizations, to provide data for their own application, as well as continue to improve input data and validation procedures. Blurring the line between mapping methods and management- and user-ready data can lead to loss of confidence in an approach, particularly at large geographic scales (Tropek et al. 2014). Thus long-term benefit and usefulness of mapping products require investment into image data acquisition along with collection and curation of training and validation datasets. This is why we provide a Google Earth Engine app to view the maps created in this study, but defer download of user-ready data to official sources where investment has already occurred.

This study describes the workflow now adapted to two initiatives aimed at mapping the distribution of coral reef ecosystems at geographic and thematic scales yet to be achieved with remote sensing. The *Allen Coral Atlas* (allencoralatlas.org) project will map all the shallow water tropical coral reefs in the world using Planet Dove satellite data and derived products (Li et al. 2019), and is funding new *in situ* field data collection. Maps from this paper are available on the *Atlas*, with regions around the world coming over 2020-2021. The *Great Barrier Reef Marine Park Authority* (gbrmpa.gov.au) has funded intensive water column correction and depth retrieval from Sentinel-2 satellite data (Roelfsema et al. 2018), to create next generation geomorphic and benthic maps (including coral type), available from early 2020. Projects and investments such as these will be critical for supporting future conservation of reefs during a time of great global change.

Growth areas for next generation reef mapping

Maps of coral reefs support a wide variety of scientific investigations and management decisions, and we hope our framework and data can support improved management and protection of reef ecosystems around the world. Our analysis demonstrates that remote sensing approaches can be used to simultaneously model the distribution of both benthic and geomorphic features, and is sufficiently flexible to adapt to other classification schemes for which georeferenced training data are available. Time-series analysis is similarly possible, and could support the detection of detrimental changes in reef environments. With appropriate investment, this would allow us to approach near real-time mapping, with a time-lag influenced only by the amount of time to process and make available image data once it has been acquired.

There are other avenues of growth in the context of improving and expanding our mapping framework. We expect some of these growth areas will include: more accurate bathymetry retrieval and deeper retrieval allowing mapping to greater depths; better tide models to create image mosaics at specific tide heights; ability to handle turbid water more accurately; incorporation of 3D information into training and validation data; transfer to next generation classification models like convolutional neural networks and deep learning; and simultaneous inclusion of change detection systems.

Existing monitoring and conservation efforts across large scales almost exclusively rely on geomorphic zonation maps to infer benthic habitat information. We expect remote sensing analyses such as ours – that provide both geomorphic and benthic habitat maps – could support studies ranging from large-scale investigations on reef dynamics, ecosystem services delivery and spatial planning for fisheries management, to finer scale site selection for restoration and ecological field studies. Some exciting examples include: better identification of refugia; three-dimensional analysis of reefs systems (incorporating bathymetry and map products); tracking patterns of resilience and understanding drivers; modelling ecosystem service benefits (e.g. wave attenuation and recruitment grounds for fisheries); understanding patterns of success and failure in restoration projects; and planning new projects to maximize recovery potential and human benefits. We hope these example research and conservation directions are only a start, and look forward to the applications our mapping framework might enable into the future.

ACKNOWLEDGMENTS

This project was initiated and funded primarily through Paul G. Allen Philanthropies and Vulcan Inc. We acknowledge the late Paul Allen and Ruth Gates for their fundamental vision and drive to enable us to work together on this critical problem. Project partners providing financial, service and personnel include: Planet Inc., National Geographic, University of Queensland, Arizona State University and University of Hawai'i. Significant support has also been provided by Google Inc., Great Barrier Reef Foundation, Trimble (Ecognition). Contributors to establishing and running the project include: Vulcan Inc [James Deutsch, Lauren Kickham, Paulina Gerstner, Charlie Whiton, Kirk Larsen, Sarah Frias Torres, Kyle Rice, Janet Greenlee, Planet Inc [Andrew Zolli, Trevor McDonald, Joe Mascaro, Joe Kington], University of Queensland [Chris M. Roelfsema, Stuart R. Phinn, Emma V. Kennedy, Dan L. Harris, Mitchell B. Lyons, Nicholas J. Murray, Eva M. Kovacs, Rodney Borrego-Acevedo, Kathryn Markey, Meredith Roe, Paul Tudman]; Arizona State University [Greg Asner, Dave Knapp, Jiwei Li, Nick S. Fabina, Heather D'Angelo]; and National Geographic [Helen Fox, Brianna Bambic, Brian Free] and Great Barrier Reef Foundation [Petra Lundgren]. We thank the other people and groups (too many to name) who tirelessly enable collection and processing of field data. Two anonymous reviewers substantially improved the scope and impact of the manuscript.

References

Achanta, R., and S. Susstrunk. 2017. Superpixels and Polygons Using Simple Non-iterative Clustering. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Andrefouet, S., F. E. Muller-Karger, J. A. Robinson, C. J. Kranenburg, D. Torres-Pulliza, S. A. Spraggins, et al. 2006. *Global assessment of modern coral reef extent and diversity for regional science and management applications: a view from space*. Reef Symposium.

Baird, M., M. Mongin, F. Rizwi, L. K. Bay, N. E. Cantin, M. Soja-Woźniak, et al. 2018. A mechanistic model of coral bleaching due to temperature-mediated light-driven reactive oxygen build-up in zooxanthellae. *Ecol. Model.* **386**, 20–37.

Beaman, R. J. 2010. Project 3D-GBR: a high-resolution depth model for the Great Barrier Reef and Coral Sea. *Marine and Tropical* Sciences Research Facility (MTSRF) Project 25i1a Final Report: 13.

Bland, L. M., T. J. Regan, M. N. Dinh, R. Ferrari, D. A. Keith, R. Lester, et al. 2017. Using multiple lines of evidence to assess the risk of ecosystem collapse. *Proc. Biol. Sci.* **284**.

Blaschke, T., G. J. Hay, M. Kelly, S. Lang, P. Hofmann, E. Addink, et al. 2014. Geographic object-based image analysis - towards a new paradigm. *ISPRS J. Photogramm. Remote Sens.* **87**, 180–191.

Callaghan, D. P., J. X. Leon, and M. I. Saunders. 2015. Wave modelling as a proxy for seagrass ecological modelling: comparing fetch and process-based predictions for a bay and reef lagoon. *Estuar. Coast. Shelf Sci.* **153**, 108–120.

Callaghan, C. T., R. E. Major, M. B. Lyons, J. M. Martin, J. H. Wilshire, R. T. Kingsford, et al. 2019. Using citizen science data to define and track restoration targets in urban areas. *J. Appl. Ecol.***56**, 1998–2006.

Coops, N. C., and M. A. Wulder. 2019. Breaking the Habit (at). *Trends Ecol. Evol.* **34**, 585–587.

Foo, S. A., and G. P. Asner. 2019. Scaling Up Coral Reef Restoration Using Remote Sensing Technology. *Frontiers in Marine. Science* **6**.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, R. Moore, et al. 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27.

Hansen, M. C., S. V. Stehman, and P. V. Potapov. 2010. Quantification of global gross forest cover loss. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 8650–8655.

Hedley, J., C. Roelfsema, and S. R. Phinn. 2009. Efficient radiative transfer model inversion for remote sensing applications. *Remote Sens. Environ.* **113**, 2527–2532.

James, G., D. Witten, T. Hastie, and R. Tibshirani. 2013. Tree based methods. In: An Introduction to Statistical Learning: with Applications in R. Springer Science & Business Media.

Keith, D. A., J. P. Rodríguez, K. M. Rodríguez-Clark, E. Nicholson, K. Aapala, A. Alonso, et al. 2013. Scientific foundations for an IUCN Red List of ecosystems. *PLoS ONE* **8**, e62111.

Li, J., D. E. Knapp, S. R. Schill, C. Roelfsema, S. Phinn, M. Silman, et al. 2019. Adaptive bathymetry estimation for shallow coastal waters using planet dove satellites. *Remote Sens. Environ.* **232**, 111302.

Lyons, M. B., D. A. Keith, S. R. Phinn, T. J. Mason, J. Elith. 2018. A comparison of resampling methods for remote sensing classification and accuracy assessment. *Remote Sens. Environ.* **208**, 145–153.

Murray, N. J., D. A. Keith, L. M. Bland, R. Ferrari, M. B. Lyons, R. Lucas, et al. 2018a. The role of satellite remote sensing in structured ecosystem risk assessments. *Sci. Total Environ.* **619**, 249–257.

Murray, N. J., D. A. Keith, D. Simpson, J. H. Wilshire, R. M. Lucas. 2018b. Remap : an online remote sensing application for land cover classification and monitoring. *Methods Ecol. Evol.* **9**, 2019–2027.

Murray, N. J., S. R. Phinn, M. DeWitt, R. Ferrari, R. Johnston, M. B. Lyons, et al. 2019. The global distribution and trajectory of tidal flats. *Nature* **565**, 222–225.

Nagendra, H., R. Lucas, J. P. Honrado, R. H. Jongman, C. Tarantino, M. Adamo, et al. 2013. Remote sensing for conservation monitoring: assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecol. Ind.* **33**, 45–59.

Pekel, J.-F., A. Cottam, N. Gorelick, and A. S. Belward. 2016. High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418–422.

Purkis, S. J., A. C. R. Gleason, C. R. Purkis, A. C. Dempsey, P. G. Renaud, M. Faisal, et al. 2019. High-resolution habitat and bathymetry maps for 65,000 sq. km of Earth's remotest coral reefs. *Coral Reefs* **38**, 467–488.

Roelfsema, C., E. Kovacs, J. C. Ortiz, N. H. Wolff, D. Callaghan, M. Wettle, et al. 2018. Coral reef habitat mapping: a combination of object-based image analysis and ecological modelling. *Remote Sens. Environ.* **208**, 27–41.

Rose, R. A., D. Byler, J. R. Eastman, E. Fleishman, G. Geller, S. Goetz, et al. 2015. Ten ways remote sensing can contribute to conservation. *Conserv. Biol.* **29**, 350–359.

Sheppard, C., and S. M. Wells. 1988. *Coral reefs of the world*. UNEP/IUCN, Gland, Switzerland.

Spalding, M., M. A. Mark D Spalding, C. Ravilious, E. P. Green. 2001. *World Atlas of Coral Reefs*. Univ of California Press, Berkeley.

Stehman, S. V., and G. M. Foody. 2019. Key issues in rigorous accuracy assessment of land cover products. *Remote Sens. Environ.* **231**, 111199.

Tropek, R., O. Sedláček, J. Beck, P. Keil, Z. Musilová, I. Šímová, et al. 2014. Comment on "High-resolution global maps of 21st-century forest cover change". *Science* **344**, 981.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1: coral reef mapping class definitions.
Appendix S2: full detailed results for accuracy assessment.

